

Order in the DP! On word order and the structure of the DP

Sarah Ouwayda and Ur Shlonsky

LSA Annual Meeting, Portland, January 8–11, 2015

Novel observations show Cinque’s (2005) phrasal movement proposal makes correct predictions on the grammaticality of word orders in Lebanese Arabic (LA) DPs. Adding an adjective yields grammatical orders Cinque (2005) cannot derive. We show that assuming an additional merge position—either for demonstratives or for numerals—derives the orders without losing Cinque’s typological predictions, and we present evidence favoring an additional *numeral* position.

1. Background: Refining Greenberg’s (1976) Universal 20, Cinque (2005) and Dryer (2006) show that the typological distribution of the order of the four elements Demonstrative, Numeral, Adjective, and Noun is extremely uneven. There are 24 possible permutations of these elements, but only 5 of them constitute the dominant orders in most languages (~77% of languages). Another 9 orders are attested, but are less frequent (~22% of languages). And there are 10 word orders that are either unattested or extremely rare, accounting all together for <1% of languages.

Cinque provides a powerful syntactic proposal to account for these facts. He assumes that the base order is that in (1), and proposes a constraint on movement, **allowing only movement of the overt NP, and phrases containing the overt NP** (henceforth “Cinque’s constraint”).

1. [_{WP} Dem [_{XP} Num [_{YP} Adj [_{NP} N]]]]

In this proposal, neither head movement, nor the movement of any constituent that does not contain the NP is allowed, and different types of permitted movements are assigned different costs: Roll-up movement with pied piping (2a) is unmarked. If, as in (2b), the NP skips certain functional structures, and moves all the way up to a higher specifier without pied-piping, the movement is somewhat marked, so word orders requiring this movement will be less frequent. The movement of an XP containing the NP, but without moving the NP up to XP's specifier first, as in (3c), is more marked, so word orders requiring it are even less frequent. Finally, splitting the NP out of a moved constituent and moving it to a higher position as in (2d) is highly marked, so word orders requiring it will be infrequent.

2. a. [WP [XP [YP [NP N]₁] Adj t₁]₂ Num t₂]₃ Dem t₃] **Unmarked**
 b. [WP [NP N]₁] Dem [XP Num [YP Adj t₁]]] **Marked**
 c. [WP [YP Adj [NP N]₁]] Dem [XP Num t₁]] **Very marked**
 d. [WP [NP N]₃] Dem [YP [XP [NP N]₁] A t₁]₂ Num t₂] **Highly marked**

Word orders requiring any other type of movement, namely head movement, or movement of a phrase not containing the overt NP, are un derivable, so they are expected to be completely unattested. And typologically, that is verified (Cinque, 2005).

2. Examining intralinguistic variation: We test the ability of Cinque's proposal to predict the grammaticality of the 24 possible permutations of Demonstrative, Numeral, Adjective, and Noun in the DP in LA. As it turns out, 9 of these 24 orders are grammatical in LA. All 9 are derivable by moving only the overt NP or phrases containing it (i.e. respecting Cinque's constraint) (3).

3. i. a. N-A-Num-Dem ~~b. Dem-Num-A-N~~ c. Num-N-A-Dem (ok in LA, frequent
d. Dem-N-A-Num e. Dem-Num-N-A typologically)

* Authors: Sarah Ouwayda, University of Geneva (sarah.ouwayda@gmail.com) & Ur Shlonsky, University of Geneva (ur.shlonsky@unige.ch).

- ii. a. N-A-Dem-Num b. N-Dem-Num-A ~~c. Num-A-N-Dem~~ (ok in LA, less
~~d. Dem-A-N-Num~~ e. Dem-N-Num-A f. N-Num-A-Dem frequent typologically)
g. N-Dem-A-Num ~~h. A-N-Num-Dem~~ i. ~~A-N-Dem-Num~~

The remaining 15 possible orders are ungrammatical in LA. 10 of them require moving a phrase not containing the NP, and are therefore correctly predicted ungrammatical (4).

4. a. Num-N-Dem-A b. Dem-A-Num-N c. Num-Dem-A-N (unattested/rare <1%)
d. N-Num-Dem-A e. Num-Dem-N-A f. A-Dem-Num-N
g. A-Dem-N-Num h. Num-A-Dem-N i. A-Num-Dem-N j. A-Num-N-Dem

The other 5 orders (the ones struck-through in (3)), are derivable by Cinque, are attested typologically, but are ungrammatical in LA. In all these orders the adjective precedes the noun. We explain their ungrammaticality by assuming, following Shlonsky (2012), that DPs in agreement-rich languages must involve NP movement above agreeing modifiers. Since adjectives agree in LA, they cannot precede the noun. So these orders cannot be grammatical.

3. Extending the Paradigm – Adding an adjective: We examine the grammaticality of the different word orders of DPs containing 5 items: a demonstrative, a numeral, an intersective adjective, a non-intersective adjective, and a noun. There are 120 possible permutations of these elements, and 117 of them are not problematic: 80 orders are ungrammatical and are ruled out for having at least one adjective preceding the noun (Shlonsky 2012). Another 9 orders are highly marked, and are ruled out because the non-intersective adjective is farther from the noun than the intersective. And another 14 orders are ungrammatical and are ruled out by Cinque’s constraint because they require moving phrases not containing the overt NP (5).

5. a. ^{*/?} Dem-N-A_i-Num-A_{ni} b. *Num-Dem-N-A_i-A_{ni} c. * Num-Dem-N-A_{ni}-A_i
d. * Num-N-Dem-A_i-A_{ni} e. * Num-N-Dem-A_{ni}-A_i f. * Num-N-A_i-Dem-A_{ni}
g. ^{*/?} N-Dem-A_i-Num-A_{ni} h. *N-Num-Dem-A_i-A_{ni} i. * N-Num-Dem-A_{ni}-A_i
j. *N-Num-A_i-Dem-A_{ni} k. *N-A_i-Dem-A_{ni}-Num l. *N-A_i-Dem-Num-A_{ni}
m. ^{*/?} N-A_i-Num-A_{ni}-Dem n. *N-A_i-Num-Dem-A_{ni}

Lastly, 14 orders are good and can be derived by moving only phrases containing the overt NP.


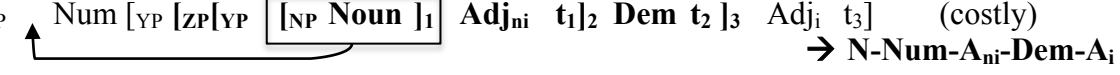
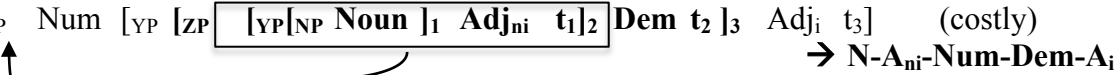
6. a. ^{ok} Dem-Num-N-A_{ni}-A_i b. ^{ok} Dem-N-Num-A_{ni}-A_i c. ^{ok} Dem-N-A_{ni}-A_i-Num
d. ^{ok} Dem-N-A_{ni}-Num-A_i e. ^{ok} Num-N-A_{ni}-A_i-Dem f. ^{ok/?} N-Dem-A_{ni}-Num-A_i
g. ^{ok/?} N-Dem-A_{ni}-A_i-Num h. ^{ok/?} N-Dem-Num-A_{ni}-A_i i. ^{ok} N-Num-A_{ni}-A_i-Dem
j. ^{ok} N-A_{ni}-A_i-Num-Dem k. ^{ok/?} N-A_{ni}-A_i-Dem-Num l. ^{ok/?} N-A_{ni}-Dem-Num-A_i
m. ^{ok/?} N-A_{ni}-Dem-A_i-Num n. ^{ok} N-A_{ni}-Num-A_i-Dem

Puzzle: There are, however, **three problematic word orders** (7) that are grammatical in LA, but that *Cinque’s Constraint* rules out. Crucially, the adjective following the demonstrative is not a reduced relative as it can also be non-intersective (“low” in terms of Cinque (2010)).

7. a. ^{ok} t-tlat mhandsiin l-madaniyyiin hol s-seeb?iin
the-three engineers the-civil these the-former (Num-N-A_{ni}-Dem-A_i)
b. ^{ok/?} l-mhandsiin t-tleeteh l-madaniyyiin hol s-seeb?iin
the- engineers the-three the-civil these the-former (N-Num-A_{ni}-Dem-A_i)
c. ^{ok} l-mhandsiin l-madaniyyiin t-tleeteh hol s-seeb?iin
engineers the-civil the-three these the-former (N-A_{ni}-Num-Dem-A_i)
‘These three former civil engineers’

While it may be tempting to abandon Cinque's Generalization based on these problematic word orders, doing so would overgenerate, and we would no longer be able to rule out (4)-(5).


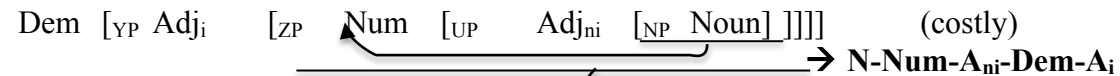
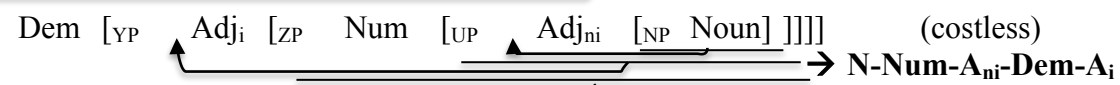
4. Proposal – Changing the base order: There are two ways to derive (7) without over-generating. The first is to assume a lower position for DEM (8), deriving (7) as in (9)-(11).

8. [WP Dem [XP Num [YP Adj [ZP Dem [Y'P Adj [NP Noun]]]]]]
9. [XP Num [YP Adj_i [ZP Dem [YP Adj_{ni} [NP Noun]]]]] (costless)
 → Num-N-A_{ni}-Dem-A_i
10. [XP Num [YP [ZP [NP Noun]₁ Adj_{ni} t₁]₂ Dem t₂]₃ Adj_i t₃] (costly)
 → N-Num-A_{ni}-Dem-A_i
11. [XP Num [YP [ZP [YP[NP Noun]₁ Adj_{ni} t₁]₂ Dem t₂]₃ Adj_i t₃] (costly)
 → N-A_{ni}-Num-Dem-A_i

Such an analysis would find its place among those proposed by Guardiano (2010), Brugè (1996), and Roberts et al. (2011), who provide cross-linguistic evidence for low DEM. It is also supported by the optional co-occurrence of a pre- and a post-nominal demonstrative in LA (12). WALS also lists 17 languages in which this co-occurrence is obligatory (Dryer et al. 2013).

12. ha l-mhandsiin hool
 DEM the-engineers DEM.pl 'These engineers' (not redundant)

Another alternative is to retain a unique position for DEM, as Cinque does, but posit a second, lower, position for *numerals* (13). The movements in (14)-(16) would derive the orders in (7).

13. [WP Dem [XP Num [YP Adj [ZP Num [Y'P Adj [NP Noun]]]]]]
14. [WP Dem [YP Adj_i [ZP Num [UP Adj_{ni} [NP Noun]]]]] (costly)
 → Num-N-A_{ni}-Dem-A_i
15. [WP Dem [YP Adj_i [ZP Num [UP Adj_{ni} [NP Noun]]]]] (costly)
 → N-Num-A_{ni}-Dem-A_i
16. [WP Dem [YP Adj_i [ZP Num [UP Adj_{ni} [NP Noun]]]]] (costless)
 → N-Num-A_{ni}-Dem-A_i

Both options therefore derive the three desirable word orders. Low DEM is supported by substantial literature. But there are reasons not to exclude a low numeral, detailed below.

4.1. REASON 1 – CASE MARKING: Standard Arabic (SA) DP-internal case patterns militate in favor of a low numeral. In the absence of numerals, case is manifested on the head noun, and adjectives concord with the noun in case (17). When a DP contains a numeral, the numeral bears the DP case, and the noun appears in the genitive (18).

17. shaahada al-mudiir-u as-saabiq-u barnaamaj-an
 watched the-boss-NOM the-former-NOM program-ACC 'the former boss watched a program'
18. shaahada thalaathat-u mudaraa-in xamsat-a barnaamij-in
 watched three-NOM boss.pl-GEN five-ACC programs-GEN '3 bosses watched 5 programs'

When an adjective is added to (18), it can match either the noun or the numeral in case (19).

19. a. waSala [thalaathat-u [[muhandisiin] saabiqiin]]
 arrived-3ms three-NOM engineer.GEN former.GEN

- b. waSala [[thalaathat-u [muhandisiin]] saabiquun]
 arrived-3ms three-NOM engineer.GEN former.NOM
 ‘Three LA former engineers arrived’

In the presence of more than one adjective, one adjective can match the noun’s case, and another adjective can match the numeral’s case. But if the adjective closest to the noun matches the numeral in case, the second adjective must too.

20. a. waSala [[thalaathat-u [[muhandisiin] saabiqin]] lubnaaniyuun]
 arrived-3ms three-NOM engineer.GEN former.GEN Lebanese-NOM
 b. *waSala thalaathat-u muhandisiinsaabiquun lubnaaniyin
 arrived-3ms three-NOM engineer.GEN former.NOM Lebanese-GEN
 ‘Three LA former engineers arrived’

Let us assume that adjectives concord in case with the case-marked complement they immediately C-command. The variation in (19a-b) can be attributed to different positions of the numeral: When the base order is [Num[Adj*[N]]], the adjective is in genitive (19a), N movement over ‘former’, followed by pied-piping over ‘Lebanese’ yields the observed order. When the base order is [Adj*[Num[N]]], the adjective bears NOM (19b), and surface order is derived by moving Num+N above the adjectives, with pied-piping. In (20a), the numeral is straddled by the adjectives, with ‘former’ immediately modifying ‘engineer’ and ‘Lebanese’ modifying ‘three former engineers’. Surface order is reached by moving N above ‘former’, then moving [_{Num}three [_{XP}[_{NP}engineer] former]] above ‘Lebanese’. Finally, the ungrammaticality of (20b) shows case is not arbitrarily assigned: An adjective concurring with the noun in GEN cannot be separated from the noun by the numeral (and adjectives concurring with the numeral). An alternative, at odds with typological generalizations, is that numerals are adjectives and can intersperse freely among different adjectives. But numerals assign case, so they cannot be adjectives.

4.2. REASON 2 – CO-OCCURRENCE: Like demonstratives, a numeral can occur in LA before or after the noun, and in (semantically odd, but grammatical) occasion, the two can co-occur (21).

21. a. t-tlatt wleed t-tleetch b. l-xams qaarraat l-arba heneh tleetch...
 the-threekids the-three the-five continents the-four t hey three...
 ‘The three kids’ (Redundant) ‘The 5 continents which are 3, are 3.’ (Contradictory)

The examples in (21) are redundant/contradictory, suggesting two numerals are indeed at play. The doubling of the demonstrative in (12), in contrast, makes no semantic contribution, suggesting that it may be a matter of concord. The obligatory co-occurrence of pre-/post-nominal demonstratives in other languages also suggests that demonstrative doubling may be concord.

The two positions for numerals also find motivations in Ouwayda’s (2011, 2014) work on transdecimal numerals, which can trigger either plural or non-plural agreement, correlating with restrictions on collective/distributive interpretations of adjectives and verbs.

5. Conclusion: While the recalcitrant orders in (7) can be handled by postulating two positions for demonstratives, our main contribution lies in proposing that there may also be two numeral positions. Note that the second Num does not adversely affect Cinque’s typological predictions: The frequent and infrequent orders are derived as he describes from Dem>Num>Adj>Noun (some are additionally derivable from Dem>Adj>Num>Noun, increasing their frequency). The rare/unattested word orders continue to be underivable or to involve very marked movements.

References:

- Brugè, Laura. (1996) Demonstrative movement in Spanish: a comparative approach. *University of Venice Working Papers in Linguistics* 6:1-53.
- Brugè, Laura. (2002) The positions of demonstratives in the extended nominal projections. In G. Cinque (ed) *Functional Structure in DP and IP*. New York/Oxford: Oxford University Press, 15-53.
- Cinque, Guglielmo. 2005. Deriving Greenberg's universal 20 and its exceptions, *Linguistic Inquiry*, v. 36:3, 315-332
- Dryer, Matthew. 2013. Order of Demonstrative and Noun. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info/chapter/88>, Accessed on 2014-06-12.)
- Guardiano, C. (2010) Demonstratives and the structure of the DP: crosslinguistic remarks. Paper presented at the Workshop on Disharmony in Nominals, Linguistics Association of Great Britain Annual Meeting, Leeds University, September 2010.
- Shlonsky, Uri. 2012. On some properties of nominals in Hebrew and Arabic, the construct state and the mechanisms of AGREE and MOVE. *Italian Journal of Linguistics*. 24.2, pp. 267-286.
- Ouwayda, Sarah. 2011. Where Plurality Is: Agreement and DP Structure. In Principal Eds. Stefan Keine and Shayne Sloggett: *Proceedings of NELS 42*. GLSA Publications: Amherst, MA.
- Ouwayda, Sarah. 2014. Where Number Lies: Plural marking, numerals, and the collective-distributive distinction, PhD. Dissertation, University of Southern California.
- Roberts, Ian. 2011. FOFC in DP: Universal 20 and the Nature of Demonstratives. ling.auf.net/lingbuzz/001502